DOCUMENT RESUME

ED 438 804 IR 019 932

AUTHOR	Choo, Chun Wei; Detlor, Brian; Turnbull, Don
TITLE	Working the Web: An Empirical Model of Web Use.
PUB DATE	1999-00-00
NOTE	10p.; Published in the Proceedings of the Hawaii International Conference on System Sciences (Maui, Hawaii, January 4-7, 2000).
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Behavior; Corporations; Foreign Countries; *Information Seeking; Internet; Online Searching; Use Studies; Users (Information); *World Wide Web
IDENTIFIERS	Computer Use

ABSTRACT

This paper investigates how knowledge workers utilize the Web to seek external information as part of their daily work. Thirty-four participants from seven companies were interviewed about their information needs and preferences. In addition, a custom-developed software application recorded each participant's Web behavior for a two week monitoring period. To understand better the significant episodes of activity identified in the Web tracking logs, a second round of personal interviews was conducted. The study found that people who use the Web as part of their work engage in four complementary modes of undirected viewing, conditioned viewing, informal search, and formal search. Moreover, each mode is characterized by information seeking moves based on recurrent sequences of browser functions and features. Based on these findings, a model of Web use is presented, as well as some suggestions for supporting and enhancing Web information seeking in organizations. (Contains 12 references.) (Author/AEF)

Reproductions supplied by EDRS are the best that can be made
from the original document.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

B. Detlor

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Working The Web: An Empirical Model of Web Use[1]

Chun Wei Choo, Brian Detlor, Don Turnbull
Faculty of Information Studies, University of Toronto, Canada
choo@fis.utoronto.ca, detlor@fis.utoronto.ca, turnbull@fis.utoronto.ca

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

1

ED 438 804

Abstract

This paper investigates how knowledge workers utilize the Web to seek external information as part of their daily work. Thirty four participants from seven companies were interviewed about their information needs and preferences. In addition, a custom-developed software application recorded each participant's Web behavior for a two week monitoring period. To understand better the significant episodes of activity identified in the Web tracking logs, a second round of personal interviews was conducted. The study found that people who use the Web as part of their work engage in four complementary modes of undirected viewing, conditioned viewing, informal search, and formal search. Moreover, each mode is characterized by information seeking moves based on recurrent sequences of browser functions and features. Based on these findings, a model of Web use is presented, as well as some suggestions for supporting and enhancing Web information seeking in organizations.

1. Introduction

The objective of this paper is to present field research on how managers and professionals use the Web as a new information technology to obtain information that helps them carry out their daily work. In terms of its potential, the Web stands to deliver a broad range of functions for content-delivery, communication, and collaboration that could benefit everyone in an organization (web pages, bulletin boards, chats, discussions, e-mail, file transfers, remote logins, and so on). However, as technology-in-use, little is known of how this tool is utilized in practice.

To address this situation, this paper describes an in-depth field investigation of organizational Web use. The paper is organized into five sections. The first familiarizes the reader with recent research on Web information seeking which relies heavily on the collection and analysis of quantitative metrics. The second provides an overview of elements from information seeking research in the organizational science and information science literature which concentrates on understanding the context behind information seeking. The third describes the paper's research design which utilizes both quantitative and qualitative measures in order to gain an understanding of both Web use metrics and the context behind organizational information seeking. The fourth presents results and findings in terms of a behavioral model of Web use. The fifth postulates, at an operational level, various ways of supporting and enhancing Web information seeking in organizations.

2. Recent Web Information Seeking Studies

Past studies have traditionally focused on the more quantitative aspects of Web information seeking activity. Catledge and Pitkow were among the first to publish a major study of Web browsing behavior by modifying the source code for a version of XMosaic, the dominant X Windows browser at the time [1]. They configured the browser to generate a client-side log file that showed user navigation strategies and interface selections. They released this modified browser to Computer Science department students who ran Mosaic from X Terminals in the various departmental computing labs at Georgia Tech. Results were measured using a task-oriented method. They determined session boundaries by analyzing the time between each event for all events, and adopted the heuristic that a lapse of 25.5 minutes or greater indicated the end of a "session." This heuristic is currently the most-commonly used for delimiting sessions. The study yielded some unexpected results. Web pages that users bookmarked did not match the most-popular sites visited as a whole from the group. Only 2% of Web pages were either saved locally or printed. These results may have been influenced by limitations in the browser (XMosaic's bookmarking capabilities), or the availability of printers in the work place. Catledge and Pitkow also hypothesized that users in their study categorized as "browsers" spend less time on a Web page than "searchers."

[1] Copyright 2000 IEEE, Published in the Proceedings of the Hawaii International Conference on System Sciences, January 4-7, 2000, Maui, Hawaii.

IR019932

BEST COPY AVAILABLE

Tauscher and Greenberg focused on the history mechanisms that Web browsers use to manage recently-requested Web pages browsed in a session of Web information seeking [2, 3]. They also used a modified XMosaic browser to collect Web browsing data for over six weeks from 23 participants. They recorded and examined the rate that Web pages were visited; how users visited old and new Web pages; the distance (in terms of URLs) between repeated Web page visits; the frequency of Web page visits, the extent of browsing in one cluster of Web pages; and repeated sequences of path-following behavior. Most significantly, they found that 58% of the pages visited during a Web browsing session were re-visits. This seems to suggest that Web information seeking may be influenced by Web browser functionality that makes it easy to go back to recently viewed pages. Tauscher and Greenberg showed that overall, users also only access a few pages frequently (60% once, and 19% twice) and browse in very small clusters of pages. They contend that Web browsing activity is a recurring system where users predominantly repeat activities they had invoked before, while still selecting new actions from the many that are possible. People explained that they revisited Web pages because the information contained by them changes; they wish to explore the page further; the page has a special purpose (e.g. search engine, home page); they are authoring a page; or the page is on a path to another revisited page. Thus, Tauscher and Greenberg identified seven Web browsing patterns: first-time visits to a cluster of pages; revisits to pages; page authoring (where the subject used Reload to view the newly modified page); use of web-based applications; hub-and-spoke visits (navigating to each new page from around a central page); a guided tour where links guide navigation through the Web pages; and a depth-first search where link paths are followed without returning to the first page in some cases.

More recently, Huberman, Pirolli, Pitkow and Lukose discovered several strong regularities of Web user surfing patterns and developed a mathematical law of surfing that determines the probability distribution of the number of pages a user visits within a Web site [4]. They started with a model of probability of the number of links a user might follow on a Web site. Next they calculated a value for the current page and related this value to the next page accessed that leads to examining the cost of continuing surfing. When the cost of moving to the next Web page is more than its expected value, the user stops Web surfing. They analyzed data collected from a sample of AOL (America Online) users for each of five days, a huge amount of data. One day alone (December 5, 1997) yielded 23,692 AOL users who collectively surfed 3,247,054 Web pages from 1,090,168 unique Web pages. This amount of data is staggering compared to previous studies of Web use.

In a related study, Huberman et. al. examined Web server logs of the Xerox external Web site in order to obtain a constrained set of Web page requests. They used "cookies" to help track the paths of individual users as they surfed through the Web site. Generally, they found a "strong fit" which was consistent through each day of the study. By applying this model along with a spreading activation algorithm, they could predict the number of requests for each Web page in a Web site. As they point out, this has implications for e-commerce applications and Web site organization, not to mention providing a more robust understanding of information seeking patterns on the Web. Overall, their study echoes other research in suggesting that surfing patterns on the Web display strong statistical regularities that can be described by a universal law. In addition, the success of the model points to the existence of utility maximizing behavior underlying surfing.

3. Information Seeking in Organizations

As described, recent research in Web use activity has focused on the more quantitative aspects of Web information seeking. We suggest that a fuller picture would emerge if quantitative analysis were complemented with qualitative analysis of the context in which Web information seeking was situated (addressing questions such as why information was needed, and how the information was used). Past research from organizational science and information science may offer help in this area. Specifically, we turn our attention to four modes of organizational scanning or external information seeking from organization science. This is followed by a description of six categories of information seeking moves derived from empirical research in information science.

3.1. Organizational Scanning Modes

The initial field work of Aguilar and the subsequent theoretical elaboration by Weick and Daft suggest that organizations scan or seek external information in four distinct modes: undirected viewing, conditioned viewing, informal search, formal search [5, 6, 7]. In undirected viewing, the individual is exposed to information with no specific informational need in mind. The overall purpose is to scan broadly in order to detect signals of change early. Many and varied sources of information are used, and large amounts of information are screened. The granularity of information is coarse, but large chunks of information are quickly dropped from attention. The goal

of broad scanning implies the use of a large number of different sources and different types of sources.

In conditioned viewing, the individual directs viewing to information about selected topics or to certain types of information. The overall purpose is to evaluate the significance of the information encountered in order to assess the general nature of the impact on the organization. The individual has isolated a number of areas of potential concern from undirected viewing, and is now sensitized to assess the significance of developments in those areas.

During informal search, the individual actively looks for information to deepen the knowledge and understanding of a specific issue. It is informal in that it involves a relatively limited and unstructured effort. The overall purpose is to gather information to elaborate an issue so as to determine the need for action by the organization.

During formal search, the individual makes a deliberate or planned effort to obtain specific information or types of information about a particular issue. Search is formal because it is structured according to some pre-established procedure or methodology. The granularity of information is fine, as search is relatively focused to find detailed information. The overall purpose is to systematically retrieve information relevant to an issue as a basis for developing a decision or course of action.

3.2. Information Seeking Behaviors

Based on their field research of information seeking by social scientists, research physicists and chemists, and engineers and research scientists in an industrial firm, Ellis et al develop a model that identified six generic categories of information seeking activities: starting, chaining, browsing, differentiating, monitoring, and extracting [8, 9, 10].

"Starting" comprises those activities that form the initial search for information—identifying sources of interest that could serve as starting points of the search. Initial sources often point to, suggest, or recommend additional references. Following up on these new leads from an initial source is the activity of "chaining". Chaining can be backward or forward. Backward chaining takes place when pointers or references from an initial source are followed, whereas forward chaining identifies and follows up on other sources that refer to an initial source. Although it can be an effective way of broadening a search, forward chaining is much less commonly used.

Having located sources and documents, "browsing" is the activity of semi-directed search in areas of potential search. The individual often simplifies browsing by looking through tables of contents, lists of titles, subject headings, names of organizations or persons, abstracts and summaries, and so on. During "differentiating", the individual filters and selects from among the sources scanned by noticing differences between the nature and quality of the information offered. The differentiation process is likely to depend on the individual's disciplinary background, prior or initial experiences with sources, word-of-mouth recommendations, or reviews in published sources.

"Monitoring" is the activity of keeping abreast of developments in an area by regularly following particular sources. The individual monitors by concentrating on a small number of what are perceived to be core sources. "Extracting" is the activity of systematically working through a particular source or sources in order to identify material of interest. Extracting tends to be labor intensive, and is more likely when there is a need for comprehensive or historical information on a topic.

Most of the information seeking behavior categories in Ellis' model are already being supported by capabilities available in common Web browser software. Thus, an individual could begin surfing the Web from one of a few favorite starting pages or sites (starting); follow links to related information resources—in both backward and forward linking directions (chaining); scan the Web pages of the sources selected (browsing); bookmark useful sources for future reference and visits (differentiating); subscribe to e-mail based services that alert the user of new information or developments (monitoring); and search a particular source or site for all information on that site on a particular topic (extracting).

4. Data Collection & Analysis

Recall that the objective of this paper to investigate and understand how managers and professionals use the Web as a new information technology to obtain information that helps them carry out their daily work. To facilitate this, we present a research methodology that combines both quantitative and qualitative measures. We believe that doing so can allow us to gain a fuller picture of organizational Web use than recent Web information seeking studies. Specifically, our research design combines quantitative analysis of Web tracking logs with qualitative analysis of personal interviews which probe the context in which Web information seeking is situated. To frame the analysis of this rich data set, we utilize both Aguilar's modes of scanning and Ellis' seeking behaviors to produce a behavioral model of organizational information seeking on the Web.

4.1. Participants

Thirty four participants from seven companies took part in the study. Since participants who regularly use the Web as part of their daily work were preferred, volunteers were canvassed through invitations at various IT-related workshops and conferences; postings at technology-focused listservs; and direct e-mail contact with colleagues and associates at large technology-oriented companies.

The seven companies comprised a major national bank, a large utility company, a national magazine publisher; a medium-sized University research library, a medium-sized marketing agency, and two small software consulting firms. The participants held jobs as IT technical specialists or analysts; managers; researchers; marketing staff; consultants; and administrative staff.

All of the users in this study primarily utilized the Web for business purposes as an integral part of their work responsibilities and activities. In most cases, participants were connected to the Internet through continuous leased-line access and used relatively high-powered machines. Many of the participants would be generally regarded as technically proficient Web users.

4.2. Data Collection

Three methods of data collection were employed: questionnaire survey; a WebTracker application that recorded Web browser actions; and personal interviews with participants.

The questionnaire survey was administered at the participants' work places, during the first site visit. The survey contained 12 questions that identified the information sources the participants used, their frequency of using these sources, and their perception of the perceived accessibility and quality of each of the sources. A wide range of sources was covered, including personal and impersonal sources (print and electronic), as well as internal and external sources. There were also questions on the amount of time and frequency of using the Web for information seeking. Furthermore, through informal conversations during the visit, research team members were able to develop a general impression of the style and scope of each participant's Web use.

The custom-developed WebTracker application was installed on each participant's computer, and it ran transparently whenever the participant's Web browser was being used. The WebTracker application was left to run on participants' computers for two-week periods. Because the WebTracker was essentially 'invisible,' it was not expected to influence participants' normal Web-use behaviors. After two weeks, WebTracker was uninstalled, and the WebTracker log file collected for analysis. The WebTracker recorded how each participant was using the browser to navigate the Web and manipulate information from the Web. Specifically, it recorded all URL calls and requests, as well as most browser menu selections, and wrote these events into a local log file on each participant's hard disk. Browser menu selections captured included "Open URL or File," "Reload," "Back," "Forward," "Add to Bookmarks," "Go to Bookmark," "Print," and "Stop." Because all URL calls and menu selections were date-time stamped as they were written into the WebTracker log, the research team was able to subsequently reconstruct move-by-move how participants looked for information on the Web during particular episodes.

The WebTracker log was pre-analyzed to prepare for personal interviews with each participant. The interview format was based on the principles of the Critical Incident Technique, in which the 'incident' to be studied should be recent, sufficiently complete, and its effects or consequences sufficiently clear [11]. In the interviews, participants described two 'critical incidents' of Web information seeking and use in reply to the following question:

"Please try to recall a recent instance in which you found important information on the Web, information that led to some significant action or decision. Would you please describe that incident for me in enough detail so that I can visualize the situation?"

Where appropriate, participants were prompted with the names of Web sites that were indicated in their WebTracker log files. Besides 'critical incidents,' participants were also invited to comment more broadly on their use of the Web, including their general Web-use strategies and preferences, as well as what they perceived to be positive and negative aspects of Web use.

4.3. Data Analysis

Data analysis was completed in two passes. The purpose of the first pass was to identify episodes of significant information seeking and to classify each episode according to the four modes of information seeking described earlier in section 3.1. The purpose of the second pass was to identify the browser activities that best characterized each episode of information seeking.

In Pass 1, significant episodes of Web-based information seeking were identified from the personal interview transcripts as well as the WebTracker logs. During interviews, participants were asked to recall "critical incidents" or significant episodes of finding and using information on the Web. By reading the transcripts, each episode was analyzed according to its

information need, amount of effort, number of web sources consulted, and information use. Based on this analysis, an episode would be categorized as one of the four modes of scanning (undirected / conditioned viewing; informal / formal searching). WebTracker logs were also examined to identify additional significant episodes. Two criteria were used to select episodes: the episode consumed a substantial amount of time and effort; or the episode was a frequently or regularly repeated activity.

Out of a total of 61 episodes identified, 12 were categorized as undirected viewing. The most common example of undirected viewing consisted of visits to general news websites such as those of NewsEdge, news.com, and newspapers. In the words of one participant, the goal was to "keep up with what's happening in the world." General news sites acted as gateways to information covering many different subject areas, and provided an efficient way of surveying current developments without a specific information need in mind. Other channels of undirected viewing included portal sites such as CANOE, and large magazine sites such as ZDnet.

Eighteen episodes were categorized as conditioned viewing. The most common examples were regular return visits to bookmarked sites, and starting from a particular page that contained links to sites of interest. Thus, a number of participants regularly visited the websites of Microsoft, Novell, and Sun Microsystems in order to monitor new content in selected sections. One participant regularly visited the Novell site for information on upcoming training courses, seminars, and software updates. Another returned to Sun's Java home page periodically to follow developments in the Electronic Commerce Framework and E-commerce tools. A third person habitually scanned the Canada Newswire Site to view press releases from the Federal and Provincial governments. Yet another customized his start-page at MSN with his own topic headings and keywords.

Twenty three episodes were categorized as informal search, and these constituted the largest group. The most common examples of informal search were when participants made use of specific query terms such as names of companies, products or technologies to perform simple searches on easily accessible search engines. There were several examples of selecting search engines that were local to a specific site (e.g. a search engine maintained by a company that only indexed its own web pages). Thus, two participants used the local search engine on the website of Forrester Research (a market research firm) to retrieve information about specific companies; another participant used the search engine at the Environmental Protection Agency to retrieve information on ventilation-heating systems for school buildings. Several of the informal searches used well-known search services from Yahoo and AltaVista.

Eight episodes were categorized as formal search. Here, participants were intending to use the information formally (e.g. to write policy or planning documents, to provide definitions). Three formal searches utilized several search engines, including meta search services. Two searches attempted to be exhaustively comprehensive: one used four meta search engines to locate a good example of an action plan that could be formally presented to a manager; the other used the DejaNews search engine to retrieve two author profiles and scan all their postings. Another search was carried out over four days, retrieving high quality resources on Women Advocacy to be included on an institutional site for International Women and Human Rights.

For each of the significant information seeking episodes categorized in Pass 1, the corresponding sections of the WebTracker log were analyzed in a second pass to determine the browser-based actions that best characterized each episode.

The WebTracker log files were tabulated into large spreadsheets with entries arranged in chronological sequence. Each entry contained a date-time value, followed by a URL or a browser menu action name. Thus it was possible to examine the information seeking moves in chronological order in each of the 61 episodes. Data about the sequence of site visits, repetitions of these sequences, movements backwards and forwards between pages, the use of bookmarking, the selection of sites from stored bookmarks, the use of search engines, printing, and other actions and events captured by the WebTracker were examined to trace the selection and development of information seeking moves over the duration of each episode. Using the criteria presented earlier based on Ellis' model in section 3.2, information seeking moves were analyzed to infer whether moves may be categorized as starting, chaining, browsing, differentiating, monitoring, or extracting.

The most common examples of starting moves took the form of participants starting their Web sessions from (1) jumpsites that contained links of interest; (2) portal sites; and (3) Intranet entry pages of their organizations. Chaining moves occurred when participants followed links from the starting page or some other page. Chaining could be in either direction (backward/forward). Browsing moves occurred when participants looked through top-level pages, examined lists of headings, or viewed sitemaps. Differentiating moves were when participants bookmarked a page, printed it, or copied its contents. Another indication of differentiating was when a person went direcly to a specific site of known content (e.g. the Microsoft site) by

entering its URL. Monitoring moves were when participants revisited favorite sites (that have for example been bookmarked or entered into a customized list/page). Although this was uncommon, another indication would be when participants signed up for email or alert services that informed them of new content on the monitored pages. Extracting moves were characterized by participants systematically working through a website to extract information of interest. A common method of extracting was to use local search engines that indexed material at their parent sites.

5. Results and Discussion

Overall, sixty-one episodes of 'significant' information seeking were identified and categorized according to their information modes. The majority of the episodes were classified as informal search (23) and conditioned viewing modes (18). A smaller number of episodes were undirected viewing (12) and formal search (8). From there, the episodes in each mode were examined in terms of their Web moves. Figure 1 illustrates how the episodes were classified across Web modes and moves in an empirical model of Web use.

In the undirected viewing episodes, data collected by the WebTracker application indicated that the most frequently occurring moves were starting and chaining. Thus, participants began at favorite starting pages (news or portal sites) and followed links that they found interesting on those pages. This was usually characterized by repetitive movement back and forth using the starting page as anchor.

One would expect to see many instances of starting and chaining in the undirected viewing mode on the Web. Starting occurs when viewers begin their web use on pre-

Information Seeking on the Web

(34 participants, 7 companies, 61 episodes)

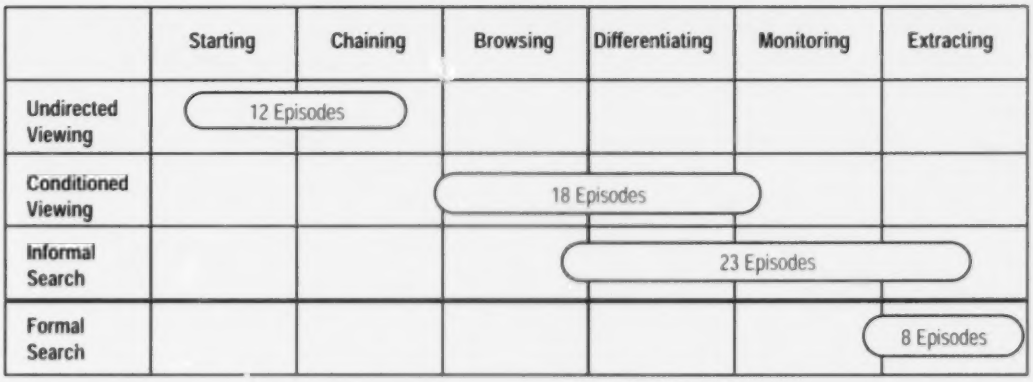

Figure 1. An Empirical Model of Web Use

selected default home pages, or when they visit a favorite page or site to begin their viewing (such as news, newspaper, or magazine sites). Chaining occurs when viewers notice items of interest (often by chance), and then follow hypertext links to more information on those items. Forward chaining of the sort just described is the most typical during undirected viewing. Backward chaining is also possible, since search engines can be used to locate other Web pages that point to the site that the user is currently at.

In the conditioned viewing episodes, the most frequently occurring moves were differentiating, browsing, and monitoring. Thus, participants selected a bookmarked page/site, or entered the URL of a site they remembered (differentiating). Another example of differentiating was when participants printed useful pages for their own files or to show to others. These sites/pages were then examined to locate new content of interest (browsing). The most important characteristic of conditioned viewing was that participants regularly or frequently returned to their selected or differentiated sites/pages to check for new information (monitoring).

One would expect browsing, differentiating, and monitoring to be common in the conditioned viewing mode on the Web. Differentiating occurs as viewers select Web sites or pages that they expect to provide

relevant information. Sites may be differentiated based on prior personal visits, or recommendations by others (such as word-of-mouth or published reviews). Differentiated sites are often bookmarked. When visiting differentiated sites, viewers browse the content by looking through tables of contents, site maps, or list of items and categories. Viewers may also monitor highly differentiated sites by returning regularly to browse, or by keeping abreast of new content (through, for example subscribing to newsletters that report new material on the site).

In the informal search episodes, the most frequently occurring moves were differentiating and localized extracting. Thus, participants went directly to selected sites where they expected that the searching they intend to do would likely yield results, e.g. going to a market research firm's site to search for company data, or to a software vendor site to search for software patches (differentiating). Searching at these sites would make use of the local search engines that were dedicated to retrieving information from those sites (localized extracting). Some participants frequently returned to specific sites to perform their informal searches (monitoring).

One would expect differentiating, extracting, and monitoring to be typical during informal search on the Web. Again, informal search is likely to be attempted at a small number of Web sites that have been differentiated by the individual, based on the individual's knowledge about these sites' information relevance, quality, affiliation, dependability, and so on. Extracting is relatively "informal" in the sense that searching would be localized to looking for information within the selected site(s). Extracting is also likely to make use of the basic, 'simple' search features or commands of the local search engine, in order to get at the most important or most recent information, without attempting to be comprehensive. Monitoring becomes more proactive if the individual sets up push channels or software agents that automatically find and deliver information based on keywords or subject headings.

In the formal search episodes, the most frequently occurring move was a relatively intensive and thorough form of extracting, compared with the localized extracting that characterized informal searching. Thus, participants systematically worked through a number of search engines or meta search engines so as to find (all) important information about a topic or item. Formal searches often involved the use of search engines known for their comprehensive coverage, currency, or the inclusion of historical data.

One would expect primarily extracting operations with some complementary monitoring activity during formal search on the Web, however for this group of participants there were no explicit instances of monitoring to support extracting. Formal search makes use of search engines that cover the Web relatively comprehensively, and that provide a powerful set of search features that can focus retrieval. Because the individual wishes not to miss any important information, there is a willingness to spend more time in the search, to learn and use complex search features, and to evaluate the sources that are found in terms of quality or accuracy. Formal search may be two-staged: multi-site searching that identifies significant sources is then followed by within-site searching. Within-site searching may involve fairly intensive foraging. Extracting may be supported by monitoring activity, again through services such as Web site alerts, push channels/agents, and e-mail announcements, in order to keep up with late-breaking information.

The distribution of information seeking episodes identified in the Web tracking data suggests that people who use the Web as part of their work engage in four complementary modes of information seeking. Each mode is set apart by its information needs, information seeking scope and effort, and the purpose of information use. Moreover, each mode of information seeking is characterized by information seeking moves that were revealed through recurrent sequences of participants' use of browser functions and features. Undirected viewing was mainly characterized by starting and chaining; conditioned viewing by differentiating, browsing, and monitoring; informal search by differentiating, and localized extracting; and formal search by systematic, thorough extracting.

The study also operationalizes and measures the six categories of information seeking behaviors identified by Ellis as browser-based actions and events [8, 9, 10]. Recurrent patterns of these actions would indicate that a user is working in a particular mode of viewing or searching on the Web. For example, repeated sequences of starting and chaining might suggest undirected viewing (moving back and forth visiting links on a starting page); while sequences of differentiating and extracting might suggest informal search (going to a bookmarked site and doing a local search).

Two other observations can be made. The first concerns monitoring which is keeping up in an area by regularly following particular core or important sources. Two forms of monitoring are possible on the Web: "pull" monitoring is when a user selects a bookmark or enters a URL to revisit a site; "push" monitoring is when a user automatically receives alerts that a monitored site has been updated. Common methods of push monitoring on the Web include subscribing to email newsletters or alerts from the monitored site; setting up a personalized profile or channel; and subscribing to services that track content

changes on selected sites. Although most participants in this study would be considered Web-savvy, very few made use of push monitoring techniques: one did use an email alert service; three others tried out a push service, but only for a limited time.

The second observation concerns extracting. On the Web, extracting is systematically searching through one or more sites in order to locate information of interest at those sites. In this study, most episodes of extracting employed basic searching strategies. For the most part, search formulations were relatively simple, with advanced features such as Boolean operators, and word truncation or proximity operators rarely utilized. This was the case even when participants appeared to be working in the formal search mode. There were no instances of participants accessing search-engine help instruction pages to improve their searches.

6. Implications

The research presented here focuses on people who use the Web to support their daily work activities. They engage in a range of complementary modes of information seeking, varying from undirected viewing that does not pursue a specific information need, to formal searching that retrieves specific information for action or decision making. Each mode of information seeking on the Web is distinguished by recurrent sequences of browser actions initiated by the information seeker. Thus, undirected viewing is characterized by starting and chaining actions; conditioned viewing is characterized by differentiating, browsing, and monitoring; informal search is characterized by differentiating and localized extracting; and formal search consists of systematic, thorough extracting.

This research has gathered empirical data to deliver a Web use model based on information seeking modes and moves. By understanding how organizational participants actually utilize the Web in work practice, we may postulate, at an operational level, ways of supporting and enhancing the information seeking of Web users as they engage in the four modes of information seeking portrayed in this study. Example suggestions are shown in Table 1 below.

Table 1. Suggestions for Enhancing Organizational Web Use

Web Use Mode	Enhancing Web Use
Undirected viewing: starting and chaining	• Introduce systems that can search or recommend new jumpsites or "similar-to" sites • Encourage people in a group to share bookmarks, Web pages, URLs • Design corporate portals to support undirected, serendipitous viewing
Conditioned viewing: browsing, differentiating, monitoring	• Train users to evaluate and escalate priority or importance of information • Make it easy for users to share Web-based information via email or in online forums • Introduce users to services that allow them to subscribe to and be notified about new content on Web pages
Informal search: differentiating, monitoring, extracting	• Pre-select high quality sources and search engines for quick, informal searches • Prepackage good search strategies developed by subject matter experts; allow users to view these strategies and learn from them • Educate users on how to evaluate information provenance and quality
Formal search: extracting	• Educate users about full range of information sources that should be considered for comprehensive search: print, online, human sources • Educate users about when to use: commercial online database services; the library o information resource center; information brokers/professionals • Train users on advanced search techniques: narrowing or broadening a search; balancing precision and recall; backward and forward chaining.

7. Acknowledgements

This research is supported by a grant from the Social Sciences and Humanities Research Council of Canada. The WebTracker application was developed by Ross

Barclay, a master's student at the Faculty of Information Studies, University of Toronto. A fuller discussion of this project and its findings is in Choo, Detlor and Turnbull [12]. The project web site is available at: http://choo.fis.utoronto.ca/esproject/.

8. References

[1] L.D. Catledge, and J.E. Pitkow, "Characterizing Browsing Strategies in the World Wide Web", 1995. *http://www.igd.fhg.de/www/www95/papers/80/userpatterns/UserPatterns.Paper4.formatted.html*

[2] L. Tauscher, and S. Greenberg, "How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems", *International Journal of Human-Computer Studies*, 47, 1997, pp. 97-137.

[3] L. Tauscher, and S. Greenburg, "Revisitation Patterns in World Wide Web Navigation", *Proceedings of CHI 97 Human Factors in Computing Systems*, S. Pemberton (ed.), Vol. 399-406.

[4] B.A. Huberman, P.L. Pirolli, J.E. Pitkow, and R.M. Lukose, "Strong regularities in World Wide Web Surfing", *Science*, 280(5360), 1998, pp. 94-97.

[5] Aguilar, F.J., *Scanning the Business Environment*, Macmillan Company, New York, 1967.

[6] K.E. Weick, and R.L. Daft, "The Effectiveness of Interpretation Systems", In *Organizational Effectiveness: A Comparison of Multiple Models*, K.S. Cameron, and D.A. Whitten (eds.), Academic Press, New York, 1983, pp. 71-93.

[7] R.L. Daft, and K.E. Weick, "Toward a Model of Organizations as Interpretation Systems", *Academy of Management Review*, 9(2), 1984, pp.284-295.

[8] D. Ellis, "A Behavioural Model for Information Retrieval System Design", *Journal of Information Science*, 15(4), 1989, pp. 237-247.

[9] D. Ellis, D. Cox, and K. Hall, "A Comparison of the Information Seeking Patterns of Researchers in the Physical and Social Sciences", *Journal of Documentation*, 49(4), 1993, pp. 356-359.

[10] D. Ellis, and M. Haugan, "Modelling the Information Seeking Patterns of Engineers and Research Scientists in an Industrial Environment", *Journal of Documentation*, 53(4), 1997, pp. 384-403.

[11] J.C. Flanagan, "The Critical Incident Technique", *Psychological Bulletin*, 51(4), 1954, pp. 327-358.

[12] Choo, C.W., Detlor, B., and Turnbull, D., *Web Work: Information Seeking and Knowledge Work on the World Wide Web*, Kluwer Academic Publishers, Dordrecht, Netherlands (2000, in press).

BEST COPY AVAILABLE